Am. J. Hum. Genet. 44:397–401, 1989

A Hypervariable Microsatellite Revealed by In Vitro Amplification of a Dinucleotide Repeat within the Cardiac Muscle Actin Gene

Michael Litt and Jeffrey A. Luty

Departments of Biochemistry and Medical Genetics, Oregon Health Sciences University, Portland

Summary

The human genome contains approximately 50,000 copies of an interspersed repeat with the sequence (dT-dG)n, where n = ~10–60. In humans, (TG)n repeats have been found in several sequenced regions. Since minisatellite regions with larger repeat elements often display extensive length polymorphisms, we suspected that (TG)n repeats ("microsatellites") might also be polymorphic. Using the polymerase chain reaction to amplify a (TG)n microsatellite in the human cardiac actin gene, we detected 12 different allelic fragments in 37 unrelated individuals, 32 of whom were heterozygous. Codominant Mendelian inheritance of fragments was observed in three families with a total of 24 children. Because of the widespread distribution of (TG)n microsatellites, polymorphisms of this type may be generally abundant and present in regions where minisatellites are rare, making such microsatellite loci very useful for linkage studies in humans.

Introduction

The human genome contains approximately 50,000 copies of an interspersed repeat with the sequence (dT-dG)n, where n = ~10–60 (Miesfeld et al. 1981; Hamada et al. 1982*b*). In humans, (TG)n repeats have been found in several sequenced regions, including the β-globin gene cluster (Miesfeld et al. 1981), the cardiac actin gene (Hamada et al. 1982*a*), and the somatostatin gene (Shen and Rutter 1984). Since minisatellite regions with larger repeat elements often display extensive length polymorphism as variable numbers of tandem repeats (VNTR) (Jeffreys et al. 1985; Nakamura et al. 1987), we suspected that (TG)n repeats ("microsatellites") might also be polymorphic. Such polymorphisms, if they exist, would probably comprise alleles differing in length by only a few base pairs and would therefore not be detectable by conventional Southern blotting. However, use of the polymerase chain reaction (PCR) (Saiki et al. 1988) with single-copy primers flanking (TG)n repeats would allow detection of microsatellite VNTRs. As a test of this idea, we chose to study the human cardiac muscle actin gene locus. This locus, previously mapped to 15q11-qter (Gunning et al. 1984), contains a (TG)n microsatellite within the fourth intron (Hamada et al. 1982*a*).

Received September 16, 1988; final revision received November 18, 1988.

Address for correspondence and reprints: Michael Litt, Department of Biochemistry, Oregon Health Sciences University, Portland, OR 97201.

0002-9297/89/4403-0013$02.00

Methods

Genomic DNAs were prepared as described elsewhere (Litt and White 1985) by using white blood cells or lymphoblast cell lines. Cell lines from Utah reference families were provided by the Coriell Institute for Medical Research. The plasmid pCM 0.9-IV, with a 900-bp *PstI* insert containing part of intron IV, exon 5, and part of intron V of the cardiac muscle actin gene, was kindly provided by Dr. T. Kakunaga.

With the published sequence as a guide (Hamada et al. 1982*a*), 20 nt primers flanking the cardiac muscle actin (TG)n repeat were synthesized on an Applied Biosystems Model 380A DNA synthesizer and purified on a 12% denaturing acrylamide gel: primer 635-TTGACCTGAATGCACTGTGA and primer 636-

TTCCATACCTGGGAACGAGT. Primer preparations were analyzed by electrophoresis on a standard DNA sequencing gel after 5′ end labeling with polynucleotide kinase and [λ^{32}P]ATP, and each was shown to contain a single fragment with <5% contamination by fragments of other than the expected size. Amplifications of 250-ng samples of genomic DNA were performed using the thermostable *Taq*I DNA polymerase and additional reagents supplied in the GeneAmp™ kit (Perkin Elmer-Cetus, Norwalk, CT). Reaction mixtures were prepared according to the instructions of the manufacturer, except that the total volume was decreased to 25 μl and 3–5 μCi of α-[^{32}P]dCTP was included in each reaction tube. Thirty-seven cycles of amplification were performed manually with denaturation for 1 min at 90°C–92°C, annealing for 2 min at 55°C, and primer extension for 2 min at 72°C. Aliquots of amplified samples were run in alternate lanes of a DNA sequencing gel. Gels were autoradiographed for 1–3 d, usually without intensifying screens. Fragment sizes were measured relative to size standards consisting of end-labeled *Sau*3A fragments of pBR322 and/or DNA sequence ladders derived from a known sequence.

For colony hybridization, 95 cosmids were randomly picked from a human genomic library in the cosmid vector Kos 1 (Maniatis et al., 1982 p. 50), kindly provided by Ed Fritsch. Poly (dG-dT) · poly (dA-dC) (Pharmacia) was nick-translated and used to probe colony filters (Grunstein and Hogness 1975) containing these cosmids. Hybridization was performed overnight at 65°C in 0.5 M sodium phosphate pH 7.0, 7% (w/v) SDS, 1% (w/v) BSA (Church and Gilbert 1984; note that this solution lacks nonradioactive DNA). Filters were washed at a maximum stringency of 0.1 × SSC/0.1% SDS at 55°C.

Results

Figure 1 shows an autoradiogram of a sequencing gel used to resolve fragments amplified from genomic DNAs from members of a 3-generation Utah family (K1329). Each lane contains one or two relatively intense "major" bands, each of which is closely associated with a cluster of up to eight less-intense minor bands spaced at 1-nt intervals. Considering that each of the major bands represents an allele, condominant Men-

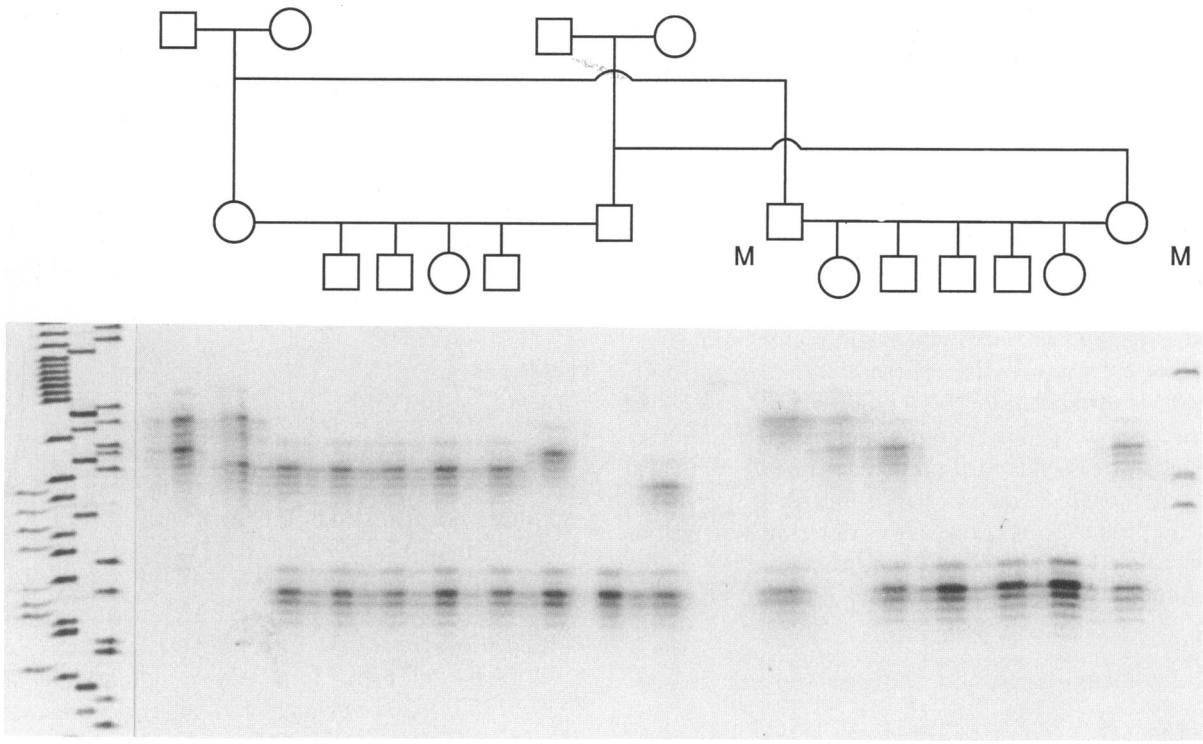

Figure 1 Autoradiograph of DNA sequencing gel illustrating inheritance of a PCR-amplified microsatellite in a portion of Utah family K1329. The samples are identified by the pedigree shown above the autoradiograph. The leftmost four lanes are DNA sequencing ladders prepared from a template of known sequence. Lanes marked M contain *Sau*3A-digested pBR322.

delian inheritance is observed in the family shown in figure 1. We have also observed codominant Mendelian inheritance of five alleles segregating in an additional 3-generation family with 15 children (data not shown). Table 1 shows the sizes and frequencies of the 12 alleles so far detected in a study of 37 unrelated Caucasian individuals. Using these allele frequencies, we calculate a PIC (Botstein et al. 1980) of .86. As expected for a VNTR polymorphism involving a dinucleotide repeat, all allelic fragments contain even numbers of nucleotides.

To study the distribution of (TG)n microsatellites in the human genome, we used poly (dG-dT) · poly (dA-dC) to probe a colony filter containing 95 cosmids (with inserts ranging from 35 to 45 kb) from a random human genomic library (fig. 2). Forty (42%) of the 95 cosmids gave strong signals. Comparison of this result with the observation (Miesfeld et al. 1981; Hamada et al. 1982*b*) that 20% of bacteriophage in a human genomic library (with inserts ranging from 12 to 22 kb; Lawn et al. 1978) hybridizes with poly (dG-dT · poly (dA-dC) confirms that (TG)n repeats are distributed very widely in the genome, with little tendency to cluster. These data also suggest that, in contrast to minisatellite VNTRs, which tend to be underrepresented in the Maniatis Charon 4A library compared with cosmid libraries (Nakamura et al. 1987), microsatellite VNTRs are equally represented in both types of genomic libraries.

The minor bands associated with allelic fragments in gels such as those of figure 1 could arise owing to impurities in the primers used for PCR. However, analysis of 5′-end-labeled primers on a DNA sequencing gel showed that fragments of sizes other than the expected 20 nt were not present in sufficient amounts to account for the generation of minor bands (data not shown). These bands may result from some artifact, such as slippage during PCR amplification, or they could reflect microheterogeneity in the lengths of (TG)n blocks in vivo. The former explanation is strongly supported by our observation that PCR amplification of the cloned cardiac actin microsatellite within the pCM 0.9-IV insert gives rise to a cluster of bands that appear very similar to the clusters seen when genomic DNA is amplified (data not shown). Since the same clone yielded an unambiguous DNA sequence (Hamada et al. 1982*a*), the variations in size which gave rise to this cluster must have occurred during the PCR amplification of the cloned DNA.

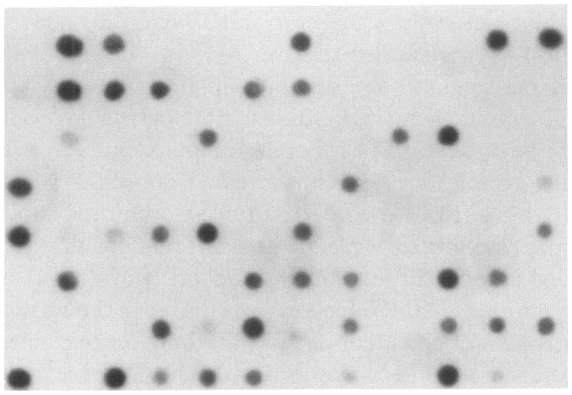

Figure 2 Colony hybridization of 95 random human genomic cosmids with poly (dG-dT) · poly (dA-dC).

Table 1

Sizes of Allelic Fragments and Their Frequencies

Size (nt)	No. Observed	Frequency
96	1	.014
92	4	.054
90	10	.135
88	7	.095
86	16	.216
84	3	.041
82	15	.203
80	1	.014
74	1	.014
72	8	.108
70	6	.081
68	2	.027

Discussion

As in the case of minisatellite VNTRs, (TG)n microsatellite polymorphisms may originate from unequal meiotic exchanges (Jeffreys et al. 1985) or from slippage during replication (Kornberg 1980). It has been reported that (TG)n repeats promote reciprocal exchange and generate unusual recombinant tetrads during yeast meiosis (Treco and Arnheim 1986).

With rare exceptions, family studies of late-onset diseases, such as Alzheimer disease, are difficult to perform because most affected individuals lack living parents and are members of small sibships. However, given a sufficient density of highly informative polymorphisms, linkage studies utilizing analytical methods such

as the affected-sib-pair (Penrose 1953; Goldin and Gershon 1988) or affected-relative (Weeks and Lange 1988) method would become generally applicable. VNTR loci are especially useful, because their full information content is available without haplotyping. In contrast, optimum exploitation of compound loci, consisting of sets of closely linked site polymorphisms, requires haplotyping, usually not feasible in small sibships that lack living parents. Unfortunately, previously described hypervariable VNTR loci tend to be biased in their distribution, with a tendency to localize in telomeric bands (Royle et al. 1987; Nakamura et al. 1988). Hence, there is an urgent need for additional hypervariable VNTRs that are broadly distributed throughout the genome. The abundance and apparently high degree of dispersion of (TG)n microsatellites and the high information content of the cardiac muscle actin microsatellite, the first such locus tested for genetic polymorphism, suggest that these loci may satisfy the need for widely distributed highly informative VNTRs. In future studies we plan to address the issues of the frequency and distribution of highly polymorphic microsatellite VNTRs in the human genome.

While the present paper was being reviewed, we became aware of a preliminary report describing similar studies by Weber and May (1988). These authors characterized 12 (TG)n microsatellites and found them all to be polymorphic, with an average PIC of .55. They also made the original observation that PCR could be used to reveal polymorphism of (TG)n repeats in the genes for insulin-like growth factor I, opsin, apolipoprotein AII, and somatostatin I (J. L. Weber, personal communication).

Acknowledgments

We thank David Kabat and John Adelman for helpful discussions and T. Kakunaga for the pCM0.9-IV probe. This work was supported by grant R01-GM32500 from the NIGMS.

References

Botstein, D., R. L. White, M. Skolnick, and R. W. Davis. 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. Am. J. Hum. Genet. **32**:314–331.

Church, G. M., and W. Gilbert. 1984. Genomic sequencing Proc. Natl. Acad. Sci. USA **81**:1991–1995.

Goldin, L. P., and E. S. Gershon. 1988. Power of the affected-sib-pair method for heterogeneous disorders. Genet. Epidemiol. **5**:35–42.

Grunstein, M., and D. S. Hogness. 1975. Colony hybridization: a method for the isolation of cloned DNAs that contain a specific gene. Proc. Natl. Acad. Sci. USA **72**:3961–3965.

Gunning, P., P. Ponte, L. Kedes, R. Eddy, and T. Shows. 1984. Chromosomal location of the co-expressed human skeletal and cardiac actin genes. Proc. Natl. Acad. Sci. USA **81**:1813–1817.

Hamada, H., M. Petrino and T. Kakunaga. 1982*a*). Molecular structure and evolutionary origin of human cardiac muscle actin gene. Proc. Natl. Acad. Sci. USA **79**:5901–5905.

———. 1982*b*. A novel repeated element with Z-DNA-forming potential is widely found in evolutionarily diverse eukaryotic genomes. Proc. Natl. Acad. Sci. USA **79**:6465–6469.

Jeffreys, A. J., V. Wilson, and S. L. Thein. 1985. Hypervariable 'minisatellite' regions in human DNA. Nature **314**: 67–73.

Kornberg, A. 1980. DNA Replication. W. H. Freeman, San Francisco.

Lawn, R. M., E. F. Fritsch, R. C. Parker, G. Blake, and T. Maniatis. 1978. The isolation and characterization of a linked δ-and β-globin gene from a cloned library of human DNA. Cell **15**:1157–1174.

Litt, M., and R. L. White. 1985. A highly polymorphic locus in human DNA revealed by cosmid derived probes. Proc. Natl. Acad. Sci. USA **82**:6206–6210.

Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Miesfeld, R., M. Krystal, and N. Arnheim. 1981. A member of a new repeated sequence family which is conserved throughout eukaryotic evolution is found between the human δ and β globin genes. Nucleic Acids Res. **9**:5931–5947.

Nakamura, Y., M. Lathrop, P. O'Connell, M. Leppert, D. Barker, E. Wright, M. Skolnick, S. Kondoleon, M. Litt, J.-M. Lalouel, and R. White. 1988. A mapped set of DNA markers for human chromosome 17. Genomics **2**:302–309.

Nakamura, Y., M. Leppert, P. O'Connell, R. Wolff, T. Holm, M. Culver, C. Martin, E. Fujimoto, M. Hoff, E. Kumlin, and R. White. 1987. Variable number of tandem repeat (VNTR) markers for human gene mapping. Science **235**: 1616–1622.

Penrose, L. S. 1953. The general purpose sib-pair linkage test. Ann. Eugenics **18**:120–124.

Royle, N. J., R. Clarkson, Z Wong, and A. J. Jeffreys. 1987. Preferential localization of hypervariable minisatellites near human telomeres. Human Gene Mapping 9 (1987). Cytogenet. Cell Genet. **46**:685.

Saiki, R. K., D. H. Gelfand, S. Stoffel, S. J. Scharf, R. Higuchi, G. T. Horn, K. B. Mullis, and H. A. Ehrlich. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science **239**:487–491.

Shen, L.-P., and W. J. Rutter. 1984. Sequence of the human somatostatin I gene. science **224**:168–170.

Treco, D., and N. Arnheim. 1986. The evolutionarily conserved repetitive sequence d(TG.AC)n promotes reciprocal exchange and generates unusual recombinant tetrads during yeast meiosis. Mol. Cell. Biol. **6**:3934–3947.

Weber, J. L., and P. E. May 1988. An abundant new class of human DNA polymorphisms. Am. J. Hum. Genet. [Suppl.] **43**:A161.

Weeks, D. E., and K. Lange. 1988. The affected-pedigree-member method of linkage analysis. Am. J. Hum. Genet. **42**:315–326.